Sparks and Deterministic Constructions of Binary Measurement Matrices from Finite Geometry

Shu-Tao Xia, Xin-Ji Liu, Yong Jiang, Hai-Tao Zheng

Abstract—**For a measurement matrix in compressed sensing, its spark (or the smallest number of columns that are linearly dependent) is an important performance parameter. The matrix with spark greater than $2k$ guarantees the exact recovery of k-sparse signals under an l_0-optimization, and the one with large spark may perform well under approximate algorithms of the l_0-optimization. Recently, Dimakis, Smarandache and Vontobel revealed the close relation between LDPC codes and compressed sensing and showed that good parity-check matrices for LDPC codes are also good measurement matrices for compressed sensing. By drawing methods and results from LDPC codes, we study the performance evaluation and constructions of binary measurement matrices in this paper. Two lower bounds of spark are obtained for general binary matrices, which improve the previously known results for real matrices in the binary case. Then, we propose two classes of deterministic binary measurement matrices based on finite geometry. Two further improved lower bounds of spark for the proposed matrices are given to show their relatively large sparks. Simulation results show that in many cases the proposed matrices perform better than Gaussian random matrices under the OMP algorithm.**

Index Terms—**Compressed sensing (CS), measurement matrix, l_0-optimization, spark, binary matrix, finite geometry, LDPC codes, deterministic construction.**

I. INTRODUCTION

Compressed sensing (CS) [1–3] is an emerging sparse sampling theory which received a large amount of attention in the area of signal processing recently. Consider a k-sparse signal $\boldsymbol{x} = (x_1, x_2, \ldots, x_n)^T \in \mathbb{R}^n$ which has at most k nonzero entries. Let $A \in \mathbb{R}^{m \times n}$ be a measurement matrix with $m \ll n$ and $\boldsymbol{y} = A\boldsymbol{x}$ be a measurement vector. Compressed sensing deals with recovering the original signal $\boldsymbol{x}$ from the measurement vector $\boldsymbol{y}$ by finding the sparsest solution to the undetermined linear system $\boldsymbol{y} = A\boldsymbol{x}$, i.e., solving the the following l_0-*optimization* problem

$$\min \|\boldsymbol{x}\|_0 \quad s.t. \quad A\boldsymbol{x} = \boldsymbol{y}, \tag{1}$$

where $\|\boldsymbol{x}\|_0 \triangleq |\{i : x_i \neq 0\}|$ denotes the l_0-norm or (Hamming) weight of $\boldsymbol{x}$. Unfortunately, it is well-known that the problem (1) is NP-hard in general [4]. In compressed sensing, there are essentially two methods to deal with it. The first one pursues greedy algorithms for l_0-optimization, such as the orthogonal matching pursuit (OMP) [5] and its modifications

[6–8]. The second method considers a convex relaxation of (1), or the l_1-*optimization* (basis pursuit) problem, as follows [2]

$$\min \|\boldsymbol{x}\|_1 \quad s.t. \quad A\boldsymbol{x} = \boldsymbol{y}, \tag{2}$$

where $\|\boldsymbol{x}\|_1 \triangleq \sum_{i=1}^{n} |x_i|$ denotes the l_1-norm of $\boldsymbol{x}$. Note that the problem (2) could be turned into a linear programming (LP) problem and thus tractable.

The construction of measurement matrix A is one of the main concerns in compressed sensing. In order to select an appropriate matrix, we need some criteria. In their earlier and fundamental works, Donoho and Elad [9] introduced the concept of *spark*. The spark of a measurement matrix A, denoted by $spark(A)$, is defined to be the smallest number of columns of A that are linearly dependent, i.e.,

$$spark(A) = \min\{\|\boldsymbol{w}\|_0 : \boldsymbol{w} \in Nullsp_{\mathbb{R}}^*(A)\}, \tag{3}$$

where

$$Nullsp_{\mathbb{R}}^*(A) \triangleq \{\boldsymbol{w} \in \mathbb{R}^n : A\boldsymbol{w} = \boldsymbol{0}, \boldsymbol{w} \neq \boldsymbol{0}\}. \tag{4}$$

Furthermore, [9] obtained several lower bounds of $spark(A)$ and showed that if

$$spark(A) > 2k, \tag{5}$$

then any k-sparse signal $\boldsymbol{x}$ can be exactly recovered by the l_0-optimization (1). In fact, we will see in the appendix that the condition (5) is also necessary for both the l_0-optimization (1) and the l_1-optimization (2). Hence, the spark is an important performance parameter of the measurement matrix. Other useful criteria include the well-known restricted isometry property (RIP) [10] and the nullspace characterization [12, 13]. Although most known constructions of measurement matrix rely on RIP, we will use spark instead in this paper since the spark is simpler and easier to deal with in some cases.

Generally, there are two main kinds of constructing methods for measurement matrices: random constructions and deterministic constructions. Many random matrices, e.g., Fourier matrices [1], Gaussian matrices, Rademacher matrices [11], *etc*, have been verified to satisfy RIP with overwhelming probability. Although random matrices perform quite well on average, there is no guarantee that a specific realization works. Moreover, storing a random matrix may require lots of storage space. On the other hand, a deterministic matrix is often generated on the fly, and some properties, e.g., spark, girth and RIP, could be verified definitely. There are many works on deterministic constructions [14–23]. Among these, constructions from coding theory [19–23] attract many attentions, e.g., Amini and Marvasti [20] used BCH codes to

construct binary, bipolar and ternary measurement matrices, and Li *et al.* [23] employed algebraic curves to generalize the constructions based on Reed-Solomon codes. In this paper, we usually use A to denote a real measurement matrix and H a binary measurement matrix.

Recently, connections between LDPC codes [24] and CS excite interest. Dimakis, Smarandache, and Vontobel [21] pointed out that the LP decoding of LDPC codes is very similar to the LP reconstruction of CS, and further showed that parity-check matrices of good LDPC codes can be used as *provably* good measurement matrices under basis pursuit. LDPC codes are a class of linear block codes, each of which is defined by the nullspace over $\mathbb{F}_2 = \{0, 1\}$ of a binary sparse $m \times n$ parity-check matrix H. Let $\mathcal{I} = \{1, 2, \ldots, n\}$ and $\mathcal{J} = \{1, 2, \ldots, m\}$ denote the sets of column indices and row indices of H, respectively. The *Tanner graph* G_H [25] corresponding to H is a bipartite graph comprising of n variable nodes labeled by the elements of $\mathcal{I}$, m check nodes labelled by the elements of $\mathcal{J}$, and the edge set $E \subseteq \{(i, j) : i \in \mathcal{I}, j \in \mathcal{J}\}$, where there is an edge $(i, j) \in E$ if and only if $h_{ji} = 1$. The *girth* g of G_H, or briefly the girth of H, is defined as the minimum length of circles in G_H. Obviously, g is always an even number and $g \geq 4$. H is said to be (γ, ρ)-*regular* if H has the uniform column weight γ and the uniform row weight ρ. The performance of an LDPC code under iterative/LP decoding over a binary erasure channel is completely determined by certain combinatorial structures, called stopping sets. A *stopping set* S of H is a subset of $\mathcal{I}$ such that the restriction of H to S, say $H(S)$, doesn't contain a row of weight one. The smallest size of a nonempty stopping set, denoted by $s(H)$, is called the stopping distance of H. Lu *et al.* [22] verified that binary sparse measurement matrices constructed by the well-known PEG algorithm [26] significantly outperform Gaussian random matrices by a series of experiments. Similar to the situation in constructing LDPC codes, matrices with girth 6 or higher are preferred in the above two works [21, 22].

In this paper, we manage to establish more connections between LDPC codes and CS. Our main contributions focus on the following two aspects.

- *Lower bounding the spark of a binary measurement matrix H.* As an important performance parameter for LDPC codes, the stopping distance $s(H)$ plays a similar role that the spark does in CS. Firstly, we show that $spark(H) \geq s(H)$, which again verifies the fact that good parity-check matrices are good measurement matrices. A special case of the lower bound is the binary corollary of the lower bound for real matrices in [9]. Then, a new general lower bound of $spark(H)$ is obtained, which improved the previous one in most cases. Furthermore, for a class of binary matrices from finite geometry, we give two further improved lower bounds to show their relatively large spark.

- *Constructing binary measurement matrices with relatively large spark.* LDPC codes based on finite geometry could be found in [27, 28]. With similar methods, two classes of deterministic constructions based on finite geometry are given, where the girth equals 4 or 6. The above lower bounds on spark ensure that the proposed matrices have relatively large spark. Simulation results show that the proposed matrices perform well and in many cases significantly better than the corresponding Gaussian random matrices under the OMP algorithm. Even in the case of girth 4, some proposed constructions still manifest good performance. Moreover, most of the proposed matrices could be put in either cyclic or quasi-cyclic form, and thus the hardware realization of sampling becomes easier and simpler.

The rest of the paper is organized as follows. In Section II we give a brief introduction to finite geometries and their parallel and quasi-cyclic structures, which result in the two classes of deterministic constructions naturally. Section III obtains our main results, or the two lower bounds of spark for general binary matrices and two further improved lower bounds for the proposed matrices from finite geometry. Simulation results and related remarks are given in Section IV. Finally, Section V concludes the paper with some discussions.

II. BINARY MATRICES FROM FINITE GEOMETRIES

Finite geometry was used to construct several classes of parity-check matrices of LDPC codes which manifest excellent performance under iterative decoding [27] [28]. We will see in the later sections that most of these structured matrices are also good measurement matrices in the sense that they often have considerably large spark and may manifest better performance than the corresponding Gaussian random matrices under the OMP algorithm. In this section, we introduce some notations and results of finite geometry [28][33, pp. 692-702].

Let $\mathbb{F}_q$ be a finite field of q elements and $\mathbb{F}_q^r$ be the r-dimensional vector space over $\mathbb{F}_q$, where $r \geq 2$. Let $EG(r, q)$ be the r-dimensional Euclidean geometry over $\mathbb{F}_q$. $EG(r, q)$ has q^r points, which are vectors of $\mathbb{F}_q^r$. The μ-flat in $EG(r, q)$ is a μ-dimensional subspace of $\mathbb{F}_q^r$ or its coset. Let $PG(r, q)$ be the r-dimensional projective geometry over $\mathbb{F}_q$. $PG(r, q)$ is defined in $\mathbb{F}_q^{r+1} \setminus \{\mathbf{0}\}$. Two nonzero vectors $\mathbf{p}, \mathbf{p}' \in \mathbb{F}_q^{r+1}$ are said to be equivalent if there is $\lambda \in \mathbb{F}_q$ such that $\mathbf{p} = \lambda \mathbf{p}'$. It is well known that all equivalence classes of $\mathbb{F}_q^{r+1} \setminus \{\mathbf{0}\}$ form points of $PG(r, q)$. $PG(r, q)$ has $(q^{r+1} - 1)/(q - 1)$ points. The μ-flat in $PG(r, q)$ is simply the set of equivalence classes contained in a $(\mu + 1)$-dimensional subspace of $\mathbb{F}_q^{r+1}$. In this paper, in order to present a unified approach, we use $FG(r, q)$ to denote either $EG(r, q)$ or $PG(r, q)$. A point is a 0-flat, a *line* is a 1-flat, and a $(r - 1)$-flat is called a *hyperplane*.

A. Incidence Matrix in Finite Geometry

For $0 \leq \mu_1 < \mu_2 \leq r$, there are $N(\mu_2, \mu_1)$ μ_1-flats contained in a given μ_2-flat and $A(\mu_2, \mu_1)$ μ_2-flats containing a given μ_1-flat, where for $EG(r, q)$ and $PG(r, q)$ respectively [28]

$$N_{EG}(\mu_2, \mu_1) = q^{\mu_2 - \mu_1} \prod_{i=1}^{\mu_1} \frac{q^{\mu_2 - i + 1} - 1}{q^{\mu_1 - i + 1} - 1}, \qquad (6)$$

$$N_{PG}(\mu_2, \mu_1) = \prod_{i=0}^{\mu_1} \frac{q^{\mu_2 - i + 1} - 1}{q^{\mu_1 - i + 1} - 1}, \qquad (7)$$

$$A_{EG}(\mu_2, \mu_1) = A_{PG}(\mu_2, \mu_1) = \prod_{i=\mu_1+1}^{\mu_2} \frac{q^{r-i+1}-1}{q^{\mu_2-i+1}-1}. \quad (8)$$

Let $n = N(r, \mu_1)$ and $J = N(r, \mu_2)$ be the numbers of μ_1-flats and μ_2-flats in $FG(r, q)$ respectively. The μ_1-flats and μ_2-flats are indexed from 1 to n and 1 to J respectively. The incidence matrix $H = (h_{ji})$ of μ_2-flat over μ_1-flat is a binary $J \times n$ matrix, where $h_{ji} = 1$ for $1 \leq j \leq J$ and $1 \leq i \leq n$ if and only if the jth μ_2-flat contains the ith μ_1-flat. The rows of H correspond to all the μ_2-flats in $FG(r, q)$ and have the same weight $N(\mu_2, \mu_1)$. The columns of H correspond to all the μ_1-flats in $FG(r, q)$ and have the same weight $A(\mu_2, \mu_1)$. Hence, H is a (γ, ρ)-regular matrix, where

$$\gamma = A(\mu_2, \mu_1), \quad \rho = N(\mu_2, \mu_1). \quad (9)$$

The incidence matrix H or H^T will be employed as measurement matrices and called respectively the *type I or type II finite geometry measurement matrix*. Moreover, by puncturing some rows or columns of H or H^T, we could construct a large amount of measurement matrices with various sizes. To obtain submatrices of H or H^T with better performance, the property of parallel structure in Euclidean Geometry are often employed as follows.

B. Parallel Structure in Euclidean Geometry

In this class of constructions, an important rule of puncturing the rows or columns of H or H^T is to make the remained submatrix as regular as possible. A possible explanation may come from Theorem 2 in the next section. This rule can be applied since the Euclidean geometry has the parallel structure and all μ_2-flats (or μ_1-flat) can be arranged by a suitable order.

Since a projective geometry does not have the parallel structure, we concentrate on $EG(r, q)$ only. Recall that a μ-flat in $EG(r, q)$ is a μ-dimensional subspace of $\mathbb{F}_q^r$ or its coset. A μ-flat contains q^μ points. Two μ-flats are either disjoint or they intersect on a flat with dimension at most $\mu - 1$. The μ-flats that correspond to the cosets of a μ-dimensional subspace of $\mathbb{F}_q^r$ (including the subspace itself) are said to be parallel to each other and form a parallel bundle. These parallel μ-flats are disjoint and contain all the points of $EG(r, q)$ with each point appearing once and only once. The number of parallel μ-flats in a parallel bundle is $q^{r-\mu}$.

There are totally $J = N_{EG}(r, \mu_2)$ μ_2-flats which consist of $K = J/q^{r-\mu_2}$ parallel bundles in $EG(r, q)$. We index these parallel bundles from $1, 2, \ldots K$. Consider the $J \times n$ incidence matrix H of μ_2-flat over μ_1-flat. All J rows of H could be divided into K bundles each of which contain $q^{r-\mu_2}$ rows, i.e., by suitable row arrangement, H could be written as

$$H = \begin{pmatrix} H_1 \\ H_2 \\ \vdots \\ H_K \end{pmatrix}, \quad (10)$$

where H_i $(1 \leq i \leq K)$ is a $q^{r-\mu_2} \times n$ submatrix of H and corresponds to the i-th parallel bundles of μ_2-flat. Clearly, the row weight of H_i remains unchanged and its column weight is 1 or 0.

Similar to the ordering of rows, the columns of H can also be ordered according to the parallel bundles in $EG(r, q)$. Hence, by deleting some row parallel bundles or column parallel bundles from H, and transposing the obtained submatrix if needed, we could construct a large amount of measurement matrices with various sizes. These will be illustrated by several examples in Section IV.

In this paper, we call H or H^T the first class of binary measurement matrices from finite geometry, and their punctured versions the second class of binary measurement matrices from finite geometry.

C. Quasi-cyclic Structure of Hyperplanes

Apart from the parallel structure of Euclidean geometry, most of the incidence matrices in Euclidean geometry and projective geometry also have cyclic or quasi-cyclic structure [28]. This is accomplished by grouping the flats of two different dimensions of a finite geometry into cyclic classes. For a Euclidean geometry, only the flats not passing through the origin are used for matrix construction. Based on this grouping of rows and columns, the incidence matrix in finite geometry consists of square submatrices (or blocks), and each of these square submatrices is a circulant matrix in which each row is a cyclic shift of the row above it and the first row is the cyclic shift of the last row. Note that by puncturing the row blocks or column blocks of the incidence matrices, the remained submatrices are often as regular as possible. In other words, this skill is compatible with the parallel structure of Euclidean geometry. Hence, the sampling process with these measurement matrices is easy and can be achieved with linear shift registers. For detailed results and discussions, we refer the readers to [28, Appendix A].

III. MAIN RESULTS

The definition of spark was introduced by Donoho and Elad [9] to help to build a theory of sparse representation that later gave birth to compressed sensing. As we see from (5), spark of the measurement matrix can be used to guarantee the exact recovery of k-sparse signals. As a result, while choosing measurement matrices, those with large sparks are prefered. However, the computation of spark is generally NP-hard [29]. In this section, we give several lower bounds of the spark for general binary matrices and the binary matrices constructed in section II by finite geometry. These theoretical results guarantee the good performance of the proposed measurement matrices under the OMP algorithm to some extent.

A. Lower Bounds for General Binary Matrices

Firstly, we give a relationship between the spark and stopping distance of a general binary matrix.

For a real vector $\boldsymbol{x} \in \mathbb{R}^n$, the support of $\boldsymbol{x}$ is defined by the set of non-zero positions, i.e., $supp(\boldsymbol{x}) \triangleq \{i : x_i \neq 0\}$. Clearly, $|supp(\boldsymbol{x})| = \|\boldsymbol{x}\|_0$.

Traditionally, a easily computable property, *coherence*, of a matrix is used to bound its spark. For a matrix $A \in \mathbb{R}^{m \times n}$

with column vectors $\boldsymbol{a}_1, \boldsymbol{a}_2, \ldots, \boldsymbol{a}_n$, the coherence $\mu(A)$ [1] is defined by [9]:

$$\mu(A) \triangleq \max_{1 \leq i \neq j \leq n} \frac{|\langle \boldsymbol{a}_i, \boldsymbol{a}_j \rangle|}{||\boldsymbol{a}_i||_2 ||\boldsymbol{a}_j||_2}, \tag{11}$$

where $\langle \boldsymbol{a}_i, \boldsymbol{a}_j \rangle \triangleq \boldsymbol{a}_i^T \boldsymbol{a}_j$ denotes the inner product of vectors. Furthermore, it is shown in [9] that

$$spark(A) \geq 1 + \frac{1}{\mu(A)}. \tag{12}$$

Note that this lower bound applies to general real matrices.

For the general binary matrix H, the next theorem shows that the spark could be lower bounded by the stopping distance.

Theorem 1: Let H be a binary matrix. Then, for any $\boldsymbol{w} \in Nullsp_{\mathbb{R}}^*(H)$, the support of $\boldsymbol{w}$ must be a stopping set of H. Moreover,

$$spark(H) \geq s(H). \tag{13}$$

Proof: Assume the contrary that $supp(\boldsymbol{w})$ is not a stopping set. By the definition of stopping set, there is one row of H containing only one '1' on $supp(\boldsymbol{w})$. Then the inner product of $\boldsymbol{w}$ and this row will be nonzero, which contradicts with the fact that $\boldsymbol{w} \in Nullsp_{\mathbb{R}}^*(A)$. Hence, $spark(A) \geq s(A)$ according to the definitions of stopping distance and spark. ∎

Remark 1: Let H be the $m \times (2^m - 1)$ parity-check matrix of a binary $[2^m - 1, m]$ Hamming code [33], which consists of all m-dimensional non-zero column vectors. It is easy to check that $spark(H) = s(H) = 3$, which implies that the lower bound (13) could be achieved. This lower bound verifies again the conclusion that good parity-check matrices are also good measurement matrices.

In particular, consider a binary $m \times n$ matrix H. Suppose the minimum column weight of H is $\gamma > 0$, and the maximum inner product of any two different columns of H is $\lambda > 0$. By (11), we have

$$\mu(H) \leq \frac{\lambda}{\gamma}.$$

Thus the lower bound (12) from [9] implies

$$spark(H) \geq 1 + \frac{\gamma}{\lambda}. \tag{14}$$

On the other hand, it was proved that $s(H) \geq 1 + \frac{\gamma}{\lambda}$ [30, 31]. Hence, the bound (14) is a natural corollary of Theorem 1.

As a matter of fact, for the general binary matrix H, we often have a tighter lower bound of its spark.

Theorem 2: Let H be a binary $m \times n$ matrix H. Suppose the minimum column weight of H is $\gamma > 0$, and the maximum inner product of any two different columns of H is $\lambda > 0$. Then

$$spark(H) \geq \frac{2\gamma}{\lambda}. \tag{15}$$

Proof: For any $\boldsymbol{w} = (w_1, w_2, \ldots, w_n)^T \in Nullsp_{\mathbb{R}}^*(H)$, we split the non-empty set $supp(\boldsymbol{w})$ into two parts $supp(\boldsymbol{w}^+)$ and $supp(\boldsymbol{w}^-)$,

$$supp(\boldsymbol{w}^+) \triangleq \{i : w_i > 0\}, \tag{16}$$
$$supp(\boldsymbol{w}^-) \triangleq \{i : w_i < 0\}. \tag{17}$$

[1]Throughout this paper, we exclude the situation that $\mu(A) = 0$ or the matrix A has a all-zero column.

Without loss of generality, we assume that $|supp(\boldsymbol{w}^+)| \geq |supp(\boldsymbol{w}^-)|$. For fixed $j \in supp(\boldsymbol{w}^+)$, by selecting the j-th column of H and all the columns in $supp(\boldsymbol{w}^-)$ of H, we get a submatrix $H(j)$. Since the column weight of H is at least γ, we could select γ rows of $H(j)$ to form a $\gamma \times (1 + |supp(\boldsymbol{w}^-)|)$ submatrix of H, say $H(\gamma, j)$, where the column corresponds to j is all 1 column. Now let's count the total number of 1's of $H(\gamma, j)$ in two ways.

- From the view of columns, since the maximum inner product of any two different columns of H is λ, each of the columns of $H(\gamma, j)$ corresponds to $supp(\boldsymbol{w}^-)$ has at most λ 1's. So the total number is at most $\gamma + \lambda |supp(\boldsymbol{w}^-)|$.
- From the view of rows, we claim that there is at least two 1's in each row of $H(\gamma, j)$, which implies the total number is at least 2γ 1's. The claim is shown as follows. Let $\boldsymbol{h}(j)$ be a row of $H(\gamma, j)$ and $\boldsymbol{h} = (h_1, \ldots, h_n)$ be its corresponding row in H. Note that $h_j = 1$. Since $\boldsymbol{w} \in Nullsp_{\mathbb{R}}^*(H)$,

$$0 = \sum_{i \in supp(\boldsymbol{w})} w_i h_i = \sum_{i \in supp(\boldsymbol{w}^+)} w_i h_i + \sum_{i \in supp(\boldsymbol{w}^-)} w_i h_i,$$

which implies that

$$- \sum_{i \in supp(\boldsymbol{w}^-)} w_i h_i = \sum_{i \in supp(\boldsymbol{w}^+)} w_i h_i \geq w_j > 0.$$

So there are at least one 1's in $\{h_i : i \in supp(\boldsymbol{w}^-)\}$ and $\boldsymbol{h}(j)$ has at least two 1's.

Therefore, $2\gamma \leq \gamma + \lambda |supp(\boldsymbol{w}^-)|$, which implies that $|supp(\boldsymbol{w}^-)| \geq \frac{\gamma}{\lambda}$. Since $|supp(\boldsymbol{w}^+)| \geq |supp(\boldsymbol{w}^-)| \geq \frac{\gamma}{\lambda}$,

$$|supp(\boldsymbol{w})| = |supp(\boldsymbol{w}^+)| + |supp(\boldsymbol{w}^-)| \geq \frac{2\gamma}{\lambda}$$

and the conclusion follows. ∎

Remark 2: Note that the matrix in Theorem 2 has girth at least 6 if $\lambda = 1$. When $\gamma \geq \lambda$, it is clear that the lower bound (15) of Theorem 2 is tighter than (14). Combining Theorem 2 with (5), we have that any k-sparse signal $\boldsymbol{x}$ can be exactly recovered by the l_0-optimization (1) if $k < \frac{\gamma}{\lambda}$.

Remark 3: Consider a complete graph on 4 vertices. The incidence matrix is

$$H = \begin{pmatrix} 1 & 1 & 1 & 0 & 0 & 0 \\ 1 & 0 & 0 & 1 & 1 & 0 \\ 0 & 1 & 0 & 1 & 0 & 1 \\ 0 & 0 & 1 & 0 & 1 & 1 \end{pmatrix}.$$

Clearly, $\gamma = 2, \lambda = 1$ and the lower bound (15) is 4. Moreover, it is easy to check that $(1, -1, 0, 0, -1, 1) \in Nullsp_{\mathbb{R}}^*(H)$ and $\{4, 5, 6\}$ is a stopping set, which implies that $spark(H) = 4$ and $s(H) = 3$. This result could be generalized to the complete graph on n nodes. Then spark and stopping distance are also 4 and 3 respectively, which implies that the lower bound (15) of Theorem 2 could be achieved and may be tighter than one in Theorem 1.

B. Lower Bounds for Binary Matrices from Finite Geometry

Clearly, the lower bound of Theorem 2 for general binary matrices applies to all ones constructed in Section II by Finite Geometry. In this subsection, we will show that for these structured matrices based on Finite Geometry, more tighter lower bound could be obtained.

Let H be the $J \times n$ incidence matrix of μ_2-flats over μ_1-flats in $FG(r,q)$, where $0 \leq \mu_1 < \mu_2 < r$, $n = N(r, \mu_1)$ and $J = N(r, \mu_2)$. Recall that H or H^T are called respectively the type-I or tyep-II finite geometry measurement matrix. The following lemma is needed to establish our results.

Lemma 1: [32] Let $0 \leq \mu_1 < \mu_2 < r$ and $1 \leq l \leq A(\mu_2, \mu_2 - 1)$. Given any l different μ_1-flats $\mathcal{F}_1, \mathcal{F}_2, \ldots, \mathcal{F}_l$ in $FG(r,q)$ and for any $1 \leq j \leq l$, there exists one $(\mu_2 - 1)$-flat $\mathcal{F}$ such that $\mathcal{F}_j \subseteq \mathcal{F}$ and $\mathcal{F}_i \not\subseteq \mathcal{F}$ for all $i = 1, \ldots, j - 1, j + 1, \ldots l$.

Theorem 3: Let r, μ_1, μ_2 be integers, $0 \leq \mu_1 < \mu_2 < r$ and H be the type-I finite geometry measurement matrix. Then

$$spark(H) \geq 2A(\mu_2, \mu_2 - 1), \qquad (18)$$

where

$$A(\mu_2, \mu_2 - 1) = \frac{q^{r-\mu_2+1} - 1}{q - 1}.$$

Proof: Let

$$u = A(\mu_2, \mu_2 - 1)$$

and assume the contrary that

$$spark(H) < 2u.$$

Select a $\boldsymbol{w} = (w_1, w_2, \ldots, w_n)^T \in Nullsp^*_{\mathbb{R}}(H)$ such that $|supp(\boldsymbol{w})| = spark(H)$. By (16) and (17), we split the nonempty set $supp(\boldsymbol{w})$ into two parts $supp(\boldsymbol{w}^+)$ and $supp(\boldsymbol{w}^-)$, and assume $|supp(\boldsymbol{w}^+)| \geq |supp(\boldsymbol{w}^-)|$ without loss of generality. Thus by the assumption

$$|supp(\boldsymbol{w}^-)| < u \quad \text{or} \quad |supp(\boldsymbol{w}^-)| \leq u - 1.$$

For fixed $j \in supp(\boldsymbol{w}^+)$, by selecting j-th column of H and all the columns in $supp(\boldsymbol{w}^-)$ of H, we get a submatrix $H(j)$. The number of columns in $H(j)$ is $1 + |supp(\boldsymbol{w}^-)|$ and not greater than u. Let $\mathcal{F}_j$ and $\{\mathcal{F}_i, i \in supp(\boldsymbol{w}^-)\}$ be the μ_1-flats corresponding to the columns of $H(j)$. By Lemma 1, there exists one $(\mu_2 - 1)$-flat $\mathcal{F}$ such that $\mathcal{F}_j \subseteq \mathcal{F}$ and $\mathcal{F}_i \not\subseteq \mathcal{F}$ for all $i \in supp(\boldsymbol{w}^-)$. There are exactly u μ_2-flats containing $\mathcal{F}$. Note that among these μ_2-flats, any two distinct μ_2-flats have no other common points except those points in $\mathcal{F}$ (see [28]). Hence, each of these u μ_2-flats contains the μ_1-flat $\mathcal{F}_j$ and for any $i \in supp(\boldsymbol{w}^-)$, there exist at most one of these u μ_2-flats containing the μ_1-flat $\mathcal{F}_i$. In other words, there exist u rows in $H(j)$ such that each of these rows has component 1 at position j and for any $i \in supp(\boldsymbol{w}^-)$, there exists at most one row that has component 1 at position i.

Let $H(u, j)$ be the $u \times (1 + |supp(\boldsymbol{w}^-)|)$ submatrix of $H(j)$ by choosing these rows, where the column corresponds to j is all 1 column. Now let's count the total number of 1's of $H(u, j)$ in two ways. The column corresponds to j has u 1's while each of the other columns has at most one 1. Thus from the view of columns, the total number of 1's in $H(u, j)$ is at most $u + |supp(\boldsymbol{w}^-)|$. On the other hand, suppose x is the number of rows in $H(u, j)$ with weight one. Then, there are $u - x$ rows with weight at least two. Thus from the view of rows, the total number of 1's in $H(u, j)$ is at least $x + 2(u - x)$. Hence, $x + 2(u - x) \leq u + |supp(\boldsymbol{w}^-)|$, which implies that $x \geq u - |supp(\boldsymbol{w}^-)| \geq 1$ by the assumption. In other words, $H(j)$ contains a row with value 1 at the position corresponding to j and 0 at other positions. Denote this row by $\boldsymbol{h}(j)$ and let $\boldsymbol{h} = (h_1, \ldots, h_n)$ be its corresponding row in H. Note that $h_j = 1$ and $h_i = 0, i \in supp(\boldsymbol{w}^-)$. Since $\boldsymbol{w} \in Nullsp^*_{\mathbb{R}}(H)$,

$$
\begin{aligned}
0 &= \sum_{i \in supp(\boldsymbol{w})} w_i h_i = \sum_{i \in supp(\boldsymbol{w}^+)} w_i h_i + \sum_{i \in supp(\boldsymbol{w}^-)} w_i h_i \\
&= \sum_{i \in supp(\boldsymbol{w}^+)} w_i h_i \geq w_j > 0,
\end{aligned}
$$

which leads to a contradiction. Therefore, the assumption is wrong and the theorem follows by (8). ∎

Remark 4: Combining Theorem 3 with (5), we have that when the type-I finite geometry measurement matrix is used, any k-sparse signal $\boldsymbol{x}$ can be exactly recovered by the l_0-optimization (1) if $k < A(\mu_2, \mu_2 - 1)$.

Remark 5: For the type-I finite geometry measurement matrix H, it is known that $s(H) \geq 1 + A(\mu_2, \mu_2 - 1)$ [32]. Thus, by Theorem 1,

$$spark(H) \geq 1 + A(\mu_2, \mu_2 - 1). \qquad (19)$$

Obviously, H has uniform column weight $\gamma = A(\mu_2, \mu_1)$. The inner product of two different columns equals to the number of μ_2-flats containing two fixed μ_1-flats. It is easy to see that the maximum inner product is $\lambda = A(\mu_2, \mu_1 + 1)$. Thus, by Theorem 2,

$$spark(H) \geq \frac{2A(\mu_2, \mu_1)}{A(\mu_2, \mu_1 + 1)}. \qquad (20)$$

It is easy to verity by (8) that

$$
\begin{aligned}
& \frac{2A(\mu_2, \mu_1)}{A(\mu_2, \mu_1 + 1)} > 1 + A(\mu_2, \mu_2 - 1) \\
\Leftrightarrow \quad & \frac{2(q^{r-\mu_1} - 1)}{q^{\mu_2 - \mu_1} - 1} > 1 + \frac{q^{r-\mu_2+1} - 1}{q - 1} \\
\Leftrightarrow \quad & q^{\mu_2 - \mu_1}(q - 2)(q^{r-\mu_2} - 1) + q^{r-\mu_2+1} - q > 0,
\end{aligned}
$$

where the last inequality always holds because of $1 \leq \mu_1 < \mu_2 < r$ and $q \geq 2$. This implies that the lower bound (20) is strictly tighter than the lower bound (19). On the other hand, the lower bound (18) of Theorem 3 is tighter than the lower bound (20). This is because

$$
\begin{aligned}
& 2A(\mu_2, \mu_2 - 1) \geq \frac{2A(\mu_2, \mu_1)}{A(\mu_2, \mu_1 + 1)} \\
\Leftrightarrow \quad & \frac{q^{r-\mu_2+1} - 1}{q - 1} \geq \frac{q^{r-\mu_1} - 1}{q^{\mu_2 - \mu_1} - 1} \\
\Leftrightarrow \quad & (q^{r-\mu_2} - 1)(q^{\mu_2 - \mu_1} - q) \geq 0.
\end{aligned}
$$

In other words, for $1 \leq \mu_1 < \mu_2 < r$, the three lower bounds (18), (20), (19) satisfies

$$2A(\mu_2, \mu_2 - 1) \geq \frac{2A(\mu_2, \mu_1)}{A(\mu_2, \mu_1 + 1)} > 1 + A(\mu_2, \mu_2 - 1), \quad (21)$$

where the inequality becomes an equality if and only if $\mu_2 = \mu_1 + 1$.

Similarly, for the type-II finite geometry measurement matrix, we obtain the following results.

Lemma 2: [32] Let $0 \le \mu_1 < \mu_2 < r$ and $1 \le l \le N(\mu_1 + 1, \mu_1)$. Given any l different μ_2-flats $\mathcal{F}_1, \mathcal{F}_2, \ldots, \mathcal{F}_l$ in $FG(r, q)$ and for any $1 \le j \le l$, there exists one $(\mu_1 + 1)$-flat $\mathcal{F}$ such that $\mathcal{F} \subseteq \mathcal{F}_j$ and $\mathcal{F} \not\subseteq \mathcal{F}_i$ for all $i = 1, \ldots, j - 1, j + 1, \ldots l$.

Theorem 4: Let r, μ_1, μ_2 be integers, $0 \le \mu_1 < \mu_2 < r$ and H^T be the type-II finite geometry measurement matrix. Then

$$spark(H^T) \ge 2N(\mu_1 + 1, \mu_1), \tag{22}$$

where for Euclidean geometry (EG) and projective geometry (PG) respectively

$$N_{EG}(\mu_1 + 1, \mu_1) = q \cdot \frac{q^{\mu_1 + 1} - 1}{q - 1},$$
$$N_{PG}(\mu_1 + 1, \mu_1) = \frac{q^{\mu_1 + 2} - 1}{q - 1}.$$

Proof: Note that the columns of H^T are rows of H. Let

$$u = N(\mu_1 + 1, \mu_1)$$

and assume the contrary that

$$spark(H^T) < 2u.$$

Select a $\boldsymbol{w} = (w_1, w_2, \ldots, w_J)^T \in Nullsp_{\mathbb{R}}^*(H^T)$ such that $|supp(\boldsymbol{w})| = spark(H^T)$. We split $supp(\boldsymbol{w})$ into $supp(\boldsymbol{w}^+)$ and $supp(\boldsymbol{w}^-)$, and assume $|supp(\boldsymbol{w}^+)| \ge |supp(\boldsymbol{w}^-)|$ without loss of generality. Thus

$$|supp(\boldsymbol{w}^-)| < u \quad \text{or} \quad |supp(\boldsymbol{w}^-)| \le u - 1.$$

For fixed $j \in supp(\boldsymbol{w}^+)$, by selecting j-th column of H^T and all the columns in $supp(\boldsymbol{w}^-)$ of H^T, we get a submatrix $H^T(j)$. The number of columns in $H^T(j)$ is $1 + |supp(\boldsymbol{w}^-)|$ and not greater than u. Let $\mathcal{F}_j$ and $\{\mathcal{F}_i, i \in supp(\boldsymbol{w}^-)\}$ be the μ_2-flats corresponding to the columns of $H^T(j)$. By Lemma 2, there exists one $(\mu_1 + 1)$-flat $\mathcal{F}$ such that $\mathcal{F} \subseteq \mathcal{F}_j$ and $\mathcal{F} \not\subseteq \mathcal{F}_i$ for all $i \in supp(\boldsymbol{w}^-)$. There are exactly u μ_1-flats contained in $\mathcal{F}$. Now, we claim that $\mathcal{F}_j$ contains these u μ_1-flats and $\mathcal{F}_i$ ($i \in supp(\boldsymbol{w}^-)$) contains at most one μ_1-flat among these u μ_1-flats. Otherwise, if $\mathcal{F}_i$ ($i \in supp(\boldsymbol{w}^-)$) contains at least two distinct μ_1-flats among these u μ_1-flats, then $\mathcal{F}_i$ must contain $\mathcal{F}$ since $\mathcal{F}$ is the only $(\mu_1 + 1)$-flat containing these two distinct μ_1-flats. This contradicts to the fact that $\mathcal{F}$ is not contained in $\mathcal{F}_i$. Hence, there exist u rows in $H^T(j)$ such that each of these rows has component 1 at position j and for any $i \in supp(\boldsymbol{w}^-)$, there exists at most one row that has component 1 at position i.

Using the same argument in the proof of Theorem 3, it leads to a contradiction. Therefore, the assumption is wrong and the theorem follows by (6) and (7). ∎

Remark 6: Combining Theorem 4 with (5), we have that when the type-II finite geometry measurement matrix is used, any k-sparse signal $\boldsymbol{x}$ can be exactly recovered by the l_0-optimization (1) if $k < N(\mu_1 + 1, \mu_1)$.

Remark 7: For the type-II finite geometry measurement matrix H^T, it is known that $s(H^T) \ge 1 + N(\mu_1 + 1, \mu_1)$ [32]. Thus, by Theorem 1,

$$spark(H^T) \ge 1 + N(\mu_1 + 1, \mu_1). \tag{23}$$

Obviously, H^T has uniform column weight $\gamma = N(\mu_2, \mu_1)$. The inner product of two different columns equals to the number of μ_1-flats contained in two fixed μ_2-flats at the same time. It is easy to see that the maximum inner product is $\lambda = N(\mu_2 - 1, \mu_1)$. Thus, by Theorem 2,

$$spark(H^T) \ge \frac{2N(\mu_2, \mu_1)}{N(\mu_2 - 1, \mu_1)}. \tag{24}$$

Using the same argument in Remark 5, we have that for $1 \le \mu_1 < \mu_2 < r$, the three lower bounds (22), (24), (23) satisfies

$$2N(\mu_1 + 1, \mu_1) \ge \frac{2N(\mu_2, \mu_1)}{N(\mu_2 - 1, \mu_1)} > 1 + N(\mu_1 + 1, \mu_1), \tag{25}$$

where the inequality becomes an equality if and only if $\mu_2 = \mu_1 + 1$ or $\mu_1 = 0$ for Euclidean geometry and $\mu_2 = \mu_1 + 1$ for projective geometry, respectively.

IV. SIMULATIONS AND ANALYSIS

In this section, we will give some simulation results on the performances for the proposed two classes of binary measurement matrices from finite geometry. The theoretical results on the sparks of these matrices in last section could explain to some extent their good performance. Afterwards, we will show by examples how to employ the parallel structure of Euclidean geometry to construct measurement matrices with flexible parameters.

All the simulations are performed under the same conditions as with [20, 23]. The upcoming figures show the percentage of perfect recovery ($SNR_{rec.} \ge 100$dB) when different sparsity orders are considered. For the generation of the k-sparse input signals, we first select the support uniformly at random and then generate the corresponding values independently by the standard normal distribution $\mathcal{N}(0, 1)$. The OMP algorithm is used to reconstruct the k-sparse input signals from the compressed measurements and the results are averaged over 5000 runs for each sparsity k. For the Gaussian random matrix each entry is chosen *i.i.d.* from $\mathcal{N}(0, 1)$. The percentages of perfect recovery of both the proposed matrix (red line) and the corresponding Gaussian random matrix (blue line) with the same size are shown in figures for comparisons.

A. Two Types of Incidence Matrices in Finite Geometry

From Theorem 3 and Theorem 4, the two types of finite geometry measurement matrices have relatively large sparks and thus we expect them to perform well under the OMP. For the type-I finite geometry measurement matrix, we expect to recover at least $(A(\mu_2, \mu_2 - 1) - 1)$-sparse signals; while for the type-II finite geometry measurement matrix, we expect to recover at least $(N(\mu_1 + 1, \mu_1) - 1)$-sparse signals.

Example 1: Let $r = 4, q = 2, \mu_2 = 3$ and $\mu_1 = 1$. The $EG(4, 2)$ consists of $J = 30$ 3-flats and $n = 120$ 1-flats. Let H be the incidence matrix of 3-flat over 1-flat in
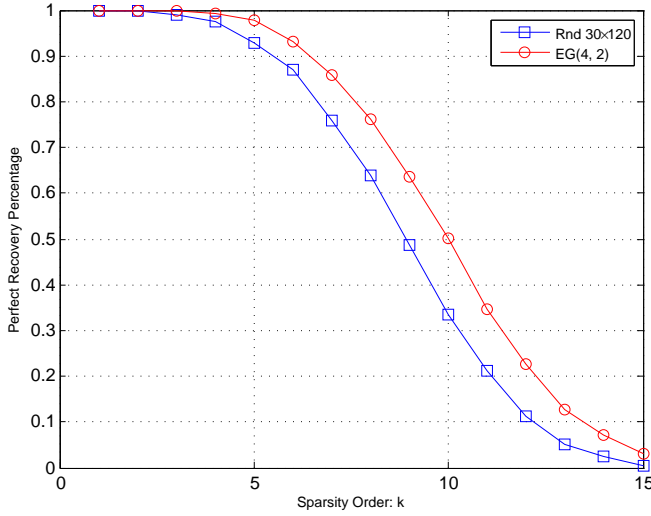
Fig. 1. Perfect recovery percentages of a type-I Euclidean geometry measurement matrix in $EG(4,2)$ with $\mu_1 = 1$, $\mu_2 = 3$ and the corresponding Gaussian random matrix.
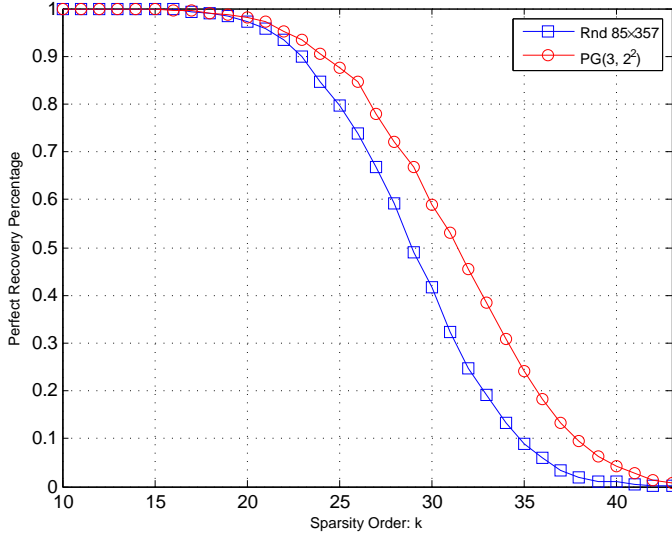

Fig. 3. Perfect recovery percentages of a type-II finite geometry measurement matrix in $EG(3,7)$ with $\mu_1 = 0$, $\mu_2 = 1$ and the corresponding Gaussian random matrix. The step size of k is 4.

Fig. 2. Perfect recovery percentages of a type-II projective geometry measurement matrix in $PG(3,2^2)$ with $\mu_1 = 0$, $\mu_2 = 1$ and the corresponding Gaussian random matrix.
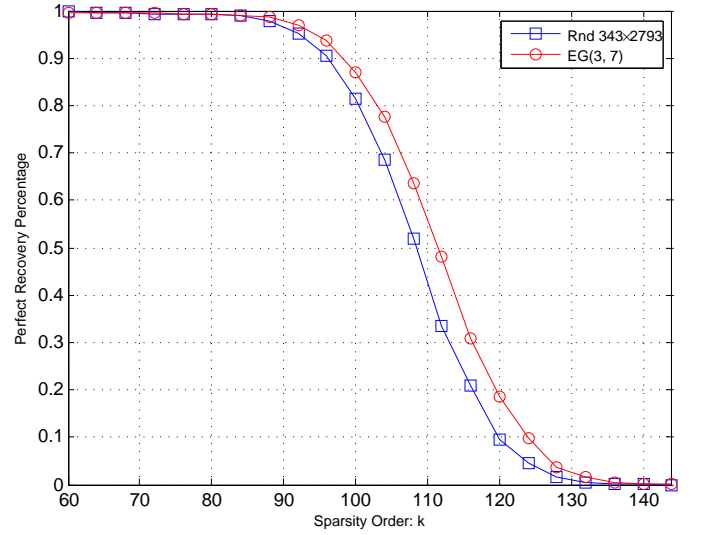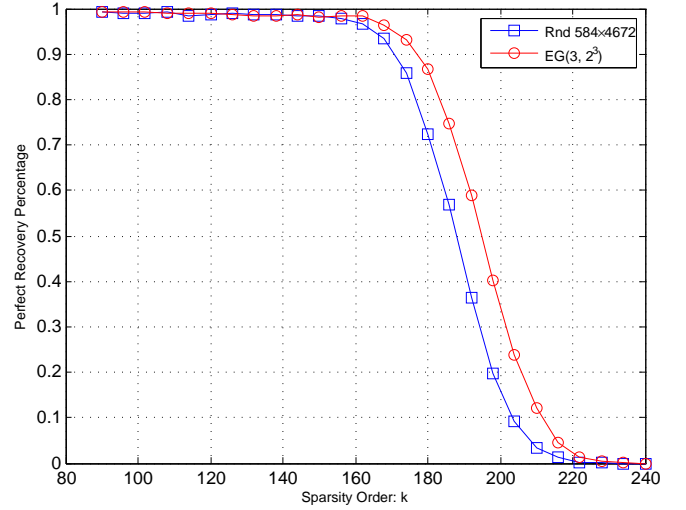

Fig. 4. Perfect recovery percentages of a type-I finite geometry measurement matrix in $EG(3,2^3)$ with $\mu_1 = 1$, $\mu_2 = 2$ and the corresponding Gaussian random matrix. The step size of k is 6.

$EG(4,2)$. Then H is a 30×120 type-I Euclidean geometry measurement matrix. H has girth 4 and is (γ, ρ)-regular, where $\gamma = A_{EG}(3,1) = 7$ and $\rho = N_{EG}(3,1) = 28$. Moreover, $spark(H) \geq 2A_{EG}(3,2) = 6$ according to Theorem 3. From Fig. 1, it is easy to find that the performance of the proposed matrix is better than that of the Gaussian random matrix. In particular, for all signals with sparsity order $k < A_{EG}(3,2) = 3$ the recovery are perfect. This example shows that some girth 4 matrices from finite geometry could also perform very well.

Example 2: Let $r = 3, q = 2^2, \mu_2 = 1$ and $\mu_1 = 0$. The $PG(3,2^2)$ consists of $J = 357$ lines and $n = 85$ points and H is the 357×85 incidence matrix of line over point in $PG(3,2^2)$. Then H^T is an 85×357 type-II projective geometry measurement matrix. H^T has girth 6 and is (γ, ρ)-regular, where $\gamma = N_{PG}(1,0) = 5$ and $\rho = A_{PG}(1,0) = 21$.

Moreover, $spark(H^T) \geq 2N_{PG}(1,0) = 10$ by Theorem 4. It is observed from Fig. 2 that H^T performs better than the Gaussian random matrix, and the sparsity order with exact recovery may exceed the one ensured by the proposed lower bound. For $k < 10$, exact recovery is obtained and the corresponding points are not plotted for clear comparisons, and the similar methods are used in the following figures.

Example 3: Let $r = 3$, $q = 7$, $\mu_2 = 1$, and $\mu_1 = 0$. The $EG(3,7)$ consists of $J = 2793$ lines and $n = 343$ points, and H is the 2793×343 incidence matrix of line over point in $EG(3,7)$. Then H^T is a type-II Euclidean geometry measurement matrix. H^T has girth 6 and is (γ, ρ)-regular, where $\gamma = N_{EG}(1,0) = 7$ and $\rho = A_{EG}(1,0) = 57$. Moreover, $spark(H) \geq 2N_{EG}(2,1) = 14$ by Theorem 4. Note that the step size of sparsity order k is 4.
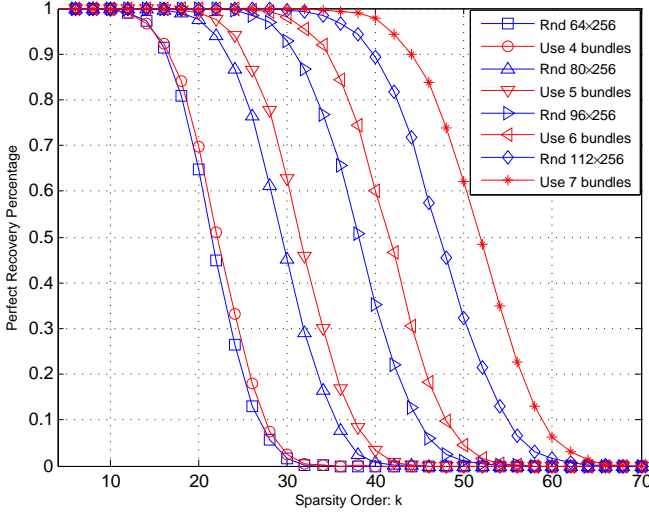
Fig. 5. Perfect recovery percentages of 4 submatrices of a type-I measurement matrix in $EG(2, 2^4)$ with $\mu_1 = 0$, $\mu_2 = 1$ and their corresponding Gaussian random matrices. The rows of the 4 submatrices from left to right are chosen according to the first 4, 5, 6 and 7 parallel bundles of lines in $EG(2, 2^4)$, respectively. The step size of k is 2.
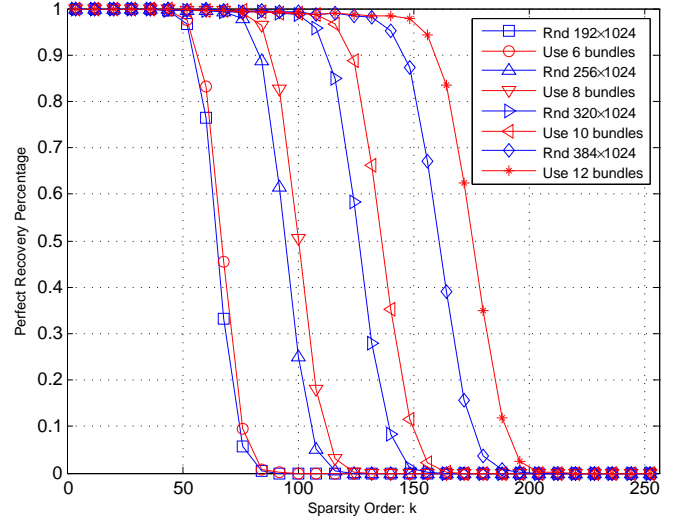

Fig. 6. Perfect recovery percentages of 4 submatrices of a type-I measurement matrix in $EG(2, 2^5)$ with $\mu_1 = 0$, $\mu_2 = 1$ and their corresponding Gaussian random matrices. The rows of the 4 submatrices from left to right are chosen according to the first 6, 8, 10 and 12 parallel bundles of lines in $EG(2, 2^5)$, respectively. The step size of k is 8.

Example 4: Let $r = 3$, $q = 2^3$, $\mu_2 = 2$, and $\mu_1 = 1$. The $EG(3, 2^3)$ consists of $J = 584$ 2-flats and $n = 4672$ 1-flats. Let H be the 584×4672 incidence matrix of 2-flat over 1-flat in $EG(3, 2^3)$. Then H is a type-I Euclidean geometry measurement matrix. H has girth 6 and is (γ, ρ)-regular, where $\gamma = A_{EG}(2, 1) = 9$ and $\rho = N_{EG}(2, 1) = 72$. Moreover, $spark(H) \geq 2A_{EG}(2, 1) = 18$ by Theorem 3. Fig. 4 shows that some matrices from finite geometry have very good performance for the moderate length of input signals (about 5000).

B. Using Parallel Structure to Obtain Measurement Matrices with Flexible Sizes

Parallel structure of Euclidean geometry is very useful to obtain various measurement matrices. Next, we show how to puncture rows or columns from the incidence matrix H or H^T by several examples.

Example 5: Let $r = 2$, $q = 2^4$, $\mu_2 = 1$ and $\mu_1 = 0$. The Euclidean plane $EG(2, 2^4)$ consists of $n = 256$ points and $J = 272$ lines. Let H be the $J \times n$ incidence matrix. Since J is close to n, both H and H^T are not suitable to be measurement matrices directly. However, according to the the parallel structure of H described in Section II, all the 272 lines can be divided into $272/16 = 17$ parallel bundles and each bundle consists of 16 lines. By (10), $H = (H_1^T, H_2^T, \ldots, H_{17}^T)^T$, where for $i = 1, \ldots, 17$, H_i consists of the 16 lines in the i-th parallel bundle. By choosing the first γ submatrices H_i, we get an $m \times n$ measurement matrix with uniform column weight γ.

Fig. 5 shows the performance of the 64×256, 80×256, 96×256, 112×256 submatrices of H which correspond to the first 4, 5, 6 and 7 parallel bundles of lines in $EG(2, 2^4)$, respectively. From Fig. 5 we can see that all of the proposed submatrices perform better than their corresponding Gaussian random matrices, and the more parallel bundles are chosen,

the better the submatrix performs, and its gain over the corresponding Gaussian random matrix becomes larger.

Example 6: Let $r = 2$, $q = 2^5$, $\mu_2 = 1$ and $\mu_1 = 0$. The Euclidean plane $EG(2, 2^5)$ consists of $n = 1024$ points and $J = 1056$ lines. Let H be the $J \times n$ incidence matrix. All the 1056 lines can be divided into $1056/32 = 33$ parallel bundles and each bundle consists of 32 lines. By (10), $H = (H_1^T, H_2^T, \ldots, H_{33}^T)^T$, where H_i consists of the 32 lines in the i-th parallel bundle. By choosing the first γ parallel bundles, we get an $m \times n$ measurement matrix with uniform column weight γ. Fig. 6 shows the performance of the 192×1024, 256×1024, 320×1024, 384×1024 submatrices of H which correspond to the first 6, 8, 10 and 12 parallel bundles of lines in $EG(2, 2^5)$, respectively. From Fig. 6 it is observed that all of the submatrices perform better than their corresponding Gaussian random matrices, and the more parallel bundles are chosen, the better the submatrix performs, and its gain over the corresponding Gaussian random matrix becomes larger.

Example 7: Consider the 320×1024 submatrix in $EG(2, 2^5)$, say H_b, in the last example. We will puncture its columns to obtain more measurement submatrices. Recall that $H = (H_1^T, H_2^T, \ldots, H_{33}^T)^T$ and the first 10 submatrices are chosen to obtain $H_b = (H_1^T, H_2^T, \ldots, H_{10}^T)^T$. For the fixed submatrix H_{11}, its corresponding 32 lines are paralleled to each other and partition the geometry. Hence, when selecting the first j lines from H_{11}, the points on these j lines are different pairwise and the total number of points is $32j$ since each line contains 32 points. By deleting the $32j$ columns corresponding to these $32j$ points from H_b, we obtain a $320 \times (1024 - 32j)$ submatrix, where is still regular.

The 4 red lines from left to right in Fig. 7 show the performance of the 320×1024, 320×896, 320×768, 320×640 submatrices of H_b which correspond that $j = 0, 4, 8, 12$, respectively. It is observed that all of the submatrices perform better than their corresponding Gaussian random matrices (the
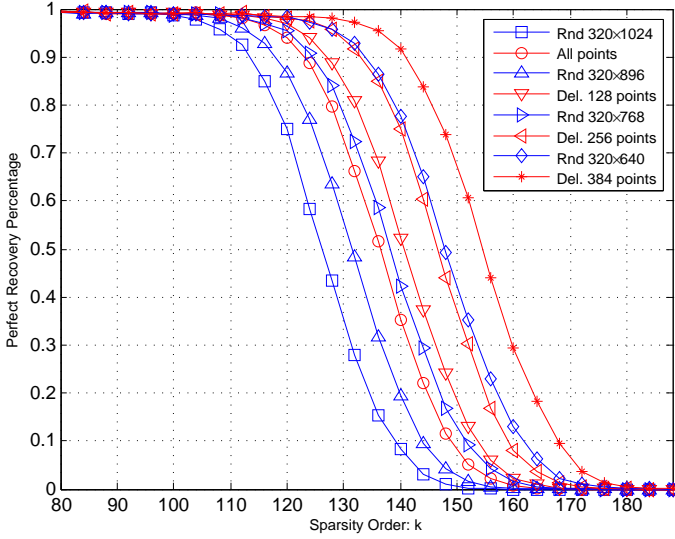
Fig. 7. Perfect recovery percentages of 4 submatrices of H_b and their corresponding Gaussian random matrices, where H_b is the 320×1024 submatrix of H in Example 6. The 4 submatrices (the red lines from left to right) are obtained by deleting 0, 128, 256, 384 columns of H_b, respectively. The step size of k is 4.

4 blue lines from left to right), but its gain becomes slightly smaller when more columns are deleted.

V. CONCLUSIONS

In this paper, by drawing methods and results from LDPC codes, we study the performance evaluation and deterministic constructions of binary measurement matrices. The spark criterion is used because its similarity to the stopping distance of an LDPC code and the fact that a matrix with large spark may perform well under the approximate algorithms of l_0-optimization, e.g., the well-known OMP algorithm. Lower bounds of spark were proposed for real matrices in [9] many years ago. When the real matrices are changed to binary matrices, better results may emerge. Firstly, two lower bounds of spark are obtained for general binary matrices, which improve the one derived from [9] in most cases. Then, we propose two classes of deterministic binary measurement matrices based on finite geometry. One class is the incidence matrix H of μ_2-flat over μ_1-flat in finite geometry $FG(r, q)$ or its transpose H^T, which are called respectively the type I or type II finite geometry measurement matrix. The other class is the submatrices of H or H^T, especially those obtained by deleting row parallel bundles or column parallel bundles from H or H^T in Euclidean geometry. In this way, we could construct a large amount of measurement matrices with various sizes. Moreover, most of the proposed matrices have cyclic or quasi-cyclic structure [28] which make the hardware realization convenient and easy. For the type I or II finite geometry measurement matrix, two further improved lower bounds of spark are given to show their relatively large spark. Finally, a lots of simulations are done according standard and comparable procedures. The simulation results show that in many cases the proposed matrices perform better than Gaussian random matrices under the OMP algorithm.

The future works may include giving more lower or upper bounds of sparks for general binary measurement matrices, determining the exact value of sparks for some classes of measurement matrices, and constructing more measurement matrices with large sparks.

APPENDIX
PROOFS OF NECESSITY

For convenience and clear statement, we write the results to the following propositions.

Proposition 1: For positive integer k, any k-sparse signal $\boldsymbol{x}$ can be exactly recovered by the l_0-optimization (1) if and only if $spark(A) > 2k$.

Proof: We only need to show the necessity. Clearly, the measurement matrix A does not have an all-0 column, which implies that $spark(A) \geq 2$. Assume the contrary that $spark(A) \leq 2k$. Select a $\boldsymbol{w} = (w_1, \ldots, w_n) \in Nullsp_{\mathbb{R}}^*(A)$ such that $\|\boldsymbol{w}\|_0 = spark(A)$. Let $a = \lfloor spark(A)/2 \rfloor$, where $\lfloor \cdot \rfloor$ is the floor function. Then

$$1 \leq a \leq k, \quad 1 \leq spark(A) - a \leq k.$$

Let b be the a-th non-zero position of $\boldsymbol{w}$ and set

$$\boldsymbol{x}' = (w_1, \ldots, w_b, 0, \ldots, 0).$$

Let $\boldsymbol{x} = \boldsymbol{x}' - \boldsymbol{w}$. Clearly,

$$\|\boldsymbol{x}'\|_0 = a, \quad \|\boldsymbol{x}\|_0 = spark(A) - a, \quad \text{and} \quad \|\boldsymbol{x}'\|_0 \leq \|\boldsymbol{x}\|_0.$$

In other words, both $\boldsymbol{x}$ and $\boldsymbol{x}'$ are k-sparse vectors, and $\boldsymbol{x}'$ may be sparser. However, since $A\boldsymbol{w} = 0$, we have that

$$A\boldsymbol{x}' = A\boldsymbol{x}' - A\boldsymbol{w} = A\boldsymbol{x},$$

which implies that $\boldsymbol{x}$ can not be exactly recovered by the l_0-optimization (1). This finishes the proof. ∎

Proposition 2: For positive integer k, any k-sparse signal $\boldsymbol{x}$ can be exactly recovered by the l_1-optimization (2) only if $spark(A) > 2k$.

Proof: It is known [12, 13, 21] that any k-sparse signal $\boldsymbol{x}$ can be exactly recovered by the l_1-optimization (2) if and only if A satisfies the so-called *nullspace property*, or for any $\boldsymbol{w} \in Nullsp_{\mathbb{R}}^*(A)$ and any $K \subseteq \{1, 2, \ldots, n\}$ with $|K| = k$,

$$\|\boldsymbol{w}_K\|_1 < \|\boldsymbol{w}_{\bar{K}}\|_1 \tag{26}$$

where $\bar{K} = \{1, 2, \ldots, n\} \setminus K$.

Assume the contrary that $spark(A) \leq 2k$. By selecting a $\boldsymbol{w} = (w_1, \ldots, w_n) \in Nullsp_{\mathbb{R}}^*(A)$ such that $\|\boldsymbol{w}\|_0 = spark(A)$, it is easy to see that $\boldsymbol{w}$ does not satisfy (26) for some k-subset K, e.g., letting K be the set of positions with the largest k $|w_i|$'s. This leads to a contradiction, which implies the conclusion. ∎

REFERENCES

[1] E. J. Candès, J. Romberg, and T. Tao, "Robust uncertainty principles: exact signal reconstruction from highly incomplete frequency information," *IEEE Trans. Inf. Theory*, vol. 52, no. 2, pp. 489–509, Feb. 2006.

[2] E. J. Candès and T. Tao, "Near-optimal signal recovery from random projections: universal encoding strategies?"

IEEE Trans. Inf. Theory, vol. 52, no. 12, pp. 5406–5425, Dec. 2006.

[3] D. L. Donoho,"Compressed sensing," *IEEE Trans. Inf. Theory*, vol. 52, no. 4, pp. 1289–1306, Apr. 2006.

[4] M. R. Garey and D. S. Johnson, *Computers and intractability: A guide to the theory of NP-completeness*, San Francisco, CA: W. H. Freeman and Company, 1979.

[5] J. Tropp and A.C. Gilbert, "Signal recovery from partial information via orthogonal matching pursuit," *IEEE Trans. Inf. Theory*, vol. 53, no. 12, pp. 4655–4666, Dec. 2007.

[6] D. Needell and J. Tropp, "CoSaMP: Iterative signal recovery from incomplete and inaccurate samples," *Appl. Comput. Harmon. Anal.,* vol. 26, no. 3, pp. 301–321, 2009.

[7] D. Needell and R. Vershynin, "Uniform uncertainty principle and signal recovery via regularized orthogonal matching pursuit," *Found. Comput. Math.,* vol. 9, no. 3, pp. 317–334, 2009.

[8] W. Dai and O. Milenkovic, "Subspace pursuit for compressive sensing signal reconstruction," *IEEE Trans. Inf. Theory*, vol. 55, no. 5, pp. 2230–2249, May 2009.

[9] D. L. Donoho and M. Elad,"Optimally sparse representation in general (nonorthogonal) dictionaries via l^1 minimization," *Proc. Nat. Acad. Sci.*, vol. 100, no. 5, pp. 2197–2202, 2003.

[10] E. J. Candès and T. Tao,"Decoding by linear programming," *IEEE Trans. Inf. Theory*, vol. 51, no. 12, pp. 4203–4215, Dec. 2005.

[11] R. Baraniuk, M. Davenport, R. DeVore and M. Wakin, "A simple proof of the restricted isometry property for random matrices," *Constr. Approx.*, vol. 28, no. 3, pp. 253–263, 2008.

[12] W. Xu and B. Hassibi, "Compressed sensing over the Grassmann manifold: A unified analytical framework," in *Proc. 46th Allerton Conf. Commun., Control, Comput.*, Monticello, IL, Sep. 2008, pp. 562–567.

[13] M. Stojnic, W. Xu, and B. Hassibi, "Compressed sensing-probabilistic analysis of a null-space characterization," in *Proc. IEEE Int. Conf. Acoust., Speech Signal Process*, LasVegas, NV, Mar.31-Apr.4, 2008, pp. 3377–3380.

[14] L. Applebauma, S. D. Howardb, S. Searlec, and R. Calderbank, "Chirp sensing codes: Deterministic compressed sensing measurements for fast recovery," *Appl. Comput. Harmon. Anal.*, vol. 26, no. 2, pp. 283–290, Mar. 2009.

[15] M. A. Iwen, "Simple deterministically constructible RIP matrices with sublinear fourier sampling requirements," in *Proc. 43rd Ann. Conf. Information Sciences and Systems* , Baltimore, MD, USA, 2009, pp. 870–875.

[16] J. Bourgain, S. Dilworth, K. Ford, S. Konyagin and D. Kutzarova, "Explicit constructions of RIP matrices and related problems," *Duke Math. J.*, vol. 159, no. 1, pp.145–185, 2011.

[17] R. A. DeVore, "Deterministic constructions of compressed sensing matrices," *J. Complexity*, vol. 23, pp. 918–925, 2007.

[18] P. Indyk, "Explicit constructions for compressed sensing

[19] S. D. Howard, A. R. Calderbank, and S. J. Searle, "A fast reconstruction algorithm for deterministic compressive sensing using second order Reed-Muller codes," in *Proc. 42nd Ann. Conf. Information Sciences and Systems* , Princeton, NJ, USA, 2008, pp. 11–15.

[20] A. Amini and F. Marvasti,"Deterministic construction of binary, bipolar, and ternary compressed sensing matrices," *IEEE Trans. Inf. Theory*, vol. 57, no. 4, pp. 2360–2370, Apr. 2011.

[21] A. G. Dimakis, R. Smarandache, and P. O. Vontobel,"LDPC codes for compressed sensing," *IEEE Trans. Inf. Theory*, vol. 58, no. 5, pp. 3093–3114, May 2012.

[22] W. Z. Lu, K. Kpalma and J. Ronisn, "Sparse binary matrices of LDPC codes for compressed sensing," in *Data Compression Conference*, Snowbird, Utah, USA, Apr. 2012, pp. 405–405.

[23] S. Li, F. Gao, G. Ge, and S. Zhang, "Deterministic construction of compressed sensing matrices via algebraic curves," *IEEE Trans. Inf. Theory*, vol. 58, no. 8, pp. 5035–5041, Apr. 2012.

[24] R. G. Gallager, "Low density parity check codes", *IRE Trans. Inf. Theory*, vol. IT-8, pp. 21–28, Jan. 1962.

[25] R. M. Tanner, "A recursive approach to low complexity codes," *IEEE Trans. Inf. Theory*, vol. 27, no. 5, pp. 533–547, Sep. 1981.

[26] X. Y. Hu, E. Eleftheriou, and D. M. Arnold, "Regular and irregular progressive edge-growth Tanner graphs," *IEEE Trans. Inf. Theory*, vol. 51, no. 1, pp. 386–398, Jan. 2005.

[27] Y. Kou, S. Lin, and M. Fossorier, "Low-density parity-check codes based on finite geometries: A rediscovery and new results," *IEEE Trans. Inf. Theory*, vol. 47, no. 7, pp. 2711–2736, Nov. 2001.

[28] H. Tang, J. Xu, S. Lin, and K. A. S. Abdel-Ghaffar, "Codes on finite geometries," *IEEE Trans. Inf. Theory*, vol. 51, no. 2, pp. 572–596, Feb. 2005.

[29] M. F. Duarte, and Y. C. Eldar,"Structured compressed sensing: From theory to applications," *IEEE Trans. Signal Processing*, vol. 59, no. 9, pp. 4053–4085, Sep. 2011.

[30] N. Kashyap and A. Vardy, "Stopping sets in codes from designs," in *Proc. IEEE Int. Symp. Inf. Theory*, Yokohama, Japan, June 29-July 4, 2003, p. 122. The full version is available online via http://www.mast.queensu.ca/∼nkashyap/Papers/stopsets.pdf.

[31] S. T. Xia and F. W. Fu, "Minimum pseudoweight and minimum pseudocodewords of LDPC codes," *IEEE Trans. Inf. Theory*, vol. 54, no. 1, pp. 480–485, Jan. 2008.

[32] S. T. Xia and F. W. Fu,"On the stopping distance of finite geometry LDPC codes," *IEEE Communications Letters*, vol. 10, no. 5, pp. 381–383, May 2006.

[33] F. J. MacWilliams and N. J. A. Sloane, *The Theory of Error-Correcting Codes*. Amsterdam, The Netherlands: North-Holland, 1981 (3rd printing).

of sparse signals," in *Proc. ACM-SIAM Symp. Discrete Algorithms*, 2008, pp. 30–33.